Claims

We claim:

1. For use with a search engine, a query pipelining system that selectively executes a user entered query made up of a string of query terms on a plurality of data sources comprising:

a query term recognizer that examines the query terms and identifies and groups terms that are intended as a phrase;

a query type recognizer that examines the query terms and categorizes the query as one of a plurality of query types;

a query intent personalizer that gathers information about the user entering the query and provides this information to the query term recognizer and query type recognizer;

a query modifier that modifies the user entered query based on the term grouping determined by the query term recognizer and the query type determined by the query type recognizer; and

a query federation module that selects data sources from the plurality of data sources and executes the modified query on the selected data sources.

2. The system of claim 1 comprising a spell checker that recognizes misspelled query terms and replaces the misspelled query terms with correctly spelled terms and wherein the query term recognizer and query type recognizer examine the correctly spelled terms.

3. The system of claim 2 wherein the spell checker augments the query with correctly spelled terms and wherein the query term recognizer and query type recognizer examine the augmented query.

4. The system of claim 1 comprising a query term list that maps possible query terms to other terms with which they are often grouped and wherein the query term recognizer refers to the query term list to identify and group terms as phrases.

5. The system of claim 1 comprising a query term list that maps possible query terms to categorization terms and wherein the query type recognizer refers to the query term list to assign a type to the query.

6. The system of claim 1 wherein the query type recognizer includes a local pattern recognizer that identifies query terms that identify the query as a local query seeking information related to a specific geographic region from which the query originated.

7. The system of claim 6 wherein the query type recognizer augments the query with information about the specific geographic region when a local query is identified.

8. The system of claim 7 wherein the query federation module selects a phone directory data source upon which to execute the local query.

9. The system of claim 1 wherein the query intent personalizer includes a context builder that retrieves information about the user entering the query.

10. The system of claim 9 wherein the information about the user includes web sites recently accessed by the user.

11. The system of claim 1 wherein the federation module includes a plurality of federation engines each dedicated to a specific data source and wherein each federation engine selectively executes the modified query on its data source based on the presence of triggering query terms in the modified query.

12. The system of claim 11 wherein each federation engine has an associated cache that saves results to previous queries that were returned to the federation engine from its data source.

13. The system of claim 1 comprising a query tracking module that records a user entered query, a modified query corresponding to the user entered query, and results returned by the modified query that were selected by the user.

14. The system of claim 4 wherein the query term list is based on user selected results to previous queries containing query terms that were grouped as a phrase in the selected results.

15. The system of claim 5 wherein the query term list maps terms to query categories based on user selected results to previous queries that included the query terms.

16. For use with a search engine query preprocessor, a method that selectively executes a query made up of a string of query terms on a plurality of data sources comprising:

obtaining context information about the origin of the query;

classifying the query as one of a set of query categories by selecting one of a set of query categories based on the presence of query terms and context information;

modifying the query to include the query category; and

executing the modified query on a data source that contains information related to the query category.

17. The method of claim 16 comprising maintaining a list of query terms mapped to categorization codes and wherein the query is classified by selecting a category code that corresponds to query terms in the query.

18. The method of claim 16 wherein one of the query categories is a local query and wherein the local queries are modified to include query context information that identifies the geographic region from which the query originates.

19. The method of claim 18 wherein the data source is a phone directory data source.

20. The method of claim 17 wherein the list of query terms is mapped to categorization codes based on a category of user selected results to previous queries containing the query terms.

21. The method of claim 17 comprising presenting the modified query to a plurality of federation engines each of which selectively executes the modified query on a data source associated with the federation engine.

22. The method of claim 21 comprising maintaining a cache for each federation engine that stores results to previously executed queries on the associated data source.

23. The method of claim 16 comprising grouping query terms as phrases and modifying the query to reflect the grouping.

24. The method of claim 23 wherein terms are grouped as phrases when previous queries containing the terms returned user selected results that grouped the terms as phrases.

25. The method of claim 16 comprising augmenting the query with correctly spelled versions of any misspelled query terms.

26. A computer readable medium comprising computer executable instructions for performing the method of claim 16.

27. For use with a search engine query preprocessor, a computer readable medium comprising computer executable instructions for executing a query made up of a string of query terms on selected data sources comprising:

obtaining context information about the user entering the query;

examining the query terms and grouping terms that are intended as a phrase based on the gathered context;

examining the query terms and categorizing the query as one of a plurality of query types based on the gathered context;

modifying the user entered query based on the term grouping and query type; and

selectively executing the modified query on the data sources.

28. The computer readable medium of claim 27 wherein the computer-executable instructions comprise replacing any misspelled query terms with correctly spelled terms examining the query containing the correctly spelled terms.

29. The computer readable medium of claim 27 wherein the computer-executable instructions comprise mapping possible query terms to other terms with which they are often grouped in a grouping list and identifying terms to be grouped as phrases based on the grouping list.

30. The computer readable medium of claim 27 wherein the computer-executable instructions comprise mapping possible query terms to categorization terms in a categorization list and assigning a type to the query based on the categorization list.

31. The computer readable medium of claim 27 wherein the computer-executable instructions comprise recognizing local query terms that identify the query as

a local query seeking information related to a specific geographic region from which the query originated.

32. The computer readable medium of claim 31 wherein the computer-executable instructions comprise augmenting the query with information about the specific geographic region when a local query is identified.

33. The computer readable medium of claim 27 wherein the computer-executable instructions comprise presenting the modified query to a plurality of federation engines each dedicated to a specific data source and wherein each federation engine selectively executes the modified query on its data source based on the presence of triggering query terms in the modified query.

34. The computer readable medium of claim 33 wherein each federation engine has an associated cache that saves results to previous queries that were returned to the federation engine from its data source.

35. For use with a search engine query preprocessor, an apparatus for executing a query made up of a string of query terms on selected data sources comprising:

means for obtaining context information about the origin of the query;

means for classifying the query as one of a set of query categories by selecting one of a set of query categories based on the presence of query terms and context information;

means for modifying the query to include the query category; and

means for executing the modified query on a data source that contains information related to the query category.

36. The apparatus of claim 35 comprising means for maintaining a list of query terms mapped to categorization codes and wherein the means for classifying selects a category code that corresponds to query terms in the query.

37. The apparatus of claim 36 wherein the list of query terms is mapped to categorization codes based on a category of results to the queries containing the query terms that were selected.

38. The apparatus of claim 35 comprising means for presenting the modified query to a plurality of federation engines each of which selectively executes the modified query on a data source associated with the federation engine.

39. The apparatus of claim 38 comprising means for maintaining a cache for each federation engine that stores results to previously executed queries on the associated data source.

40. The apparatus of claim 35 comprising means for grouping query terms as phrases and means for modifying the query to reflect the grouping.

41. The apparatus of claim 35 comprising means for augmenting the query with correctly spelled versions of any misspelled query terms.